

The “Supervenience Argument”: Kim’s Challenge to Nonreductive Physicalism

AUSONIO MARRAS

University of Western Ontario
amarras@uwo.ca

JUHANI YLI-VAKKURI

McGill University
juhani.yli-vakkuri@mail.mcgill.ca

ABSTRACT. Jaegwon Kim’s “supervenience argument” purports to show that epiphenomenalism about the mental follows from premises that any nonreductive physicalist should find acceptable. Kim regards his argument as a *reductio ad absurdum* of nonreductive physicalism. We reconstruct and evaluate the latest version of Kim’s argument. We argue that the premises of Kim’s argument are much less innocent than they may appear. In particular, we single out for criticism an unstated assumption about the identity conditions of events, and we argue that this assumption could be seen as all by itself implying that nonreductive physicalism is false, thus begging the question against that position. It is also dubious, we argue, whether Kim’s unstated assumption is even consistent with one of the stated assumptions of his argument, “the principle of causal exclusion”, given a standard understanding of causal overdetermination. We conclude with some polemical remarks about the conception of causation presupposed by Kim’s argument—a conception that appears to depart from that at work in science and commonsense discourse.

1. *Introduction*¹

Many philosophers have worried that physical causation may “exclude” mental causation—physical and mental causes “compete” for efficacy, and because of some principle (the causal closure of the physical domain, for example), physical causes inevitably “win”. This picturesque language is common, but explicit arguments, which identify *prima facie plausible* principles from which epiphenomenalism about the mental would logically

¹ We would like to thank Ian Gold for helpful comments on an earlier draft of this paper.

follow, are less so. Jaegwon Kim's "supervenience argument" represents perhaps the most influential attempt to construct such an argument. In Kim's case, the argument is presented as a *reductio ad absurdum* of *nonreductive physicalism*. The present paper is an attempt to get clear on just what Kim's supervenience argument is, and how, or whether, it works. We will only discuss Kim's most recent formulation of the supervenience argument, or arguments—Kim gives us two "versions" of the argument ("Completion 1" and "Completion 2"). This formulation is found in ch. 2 of Kim's most recent book, *Physicalism, or Something Near Enough* (Kim 2005; henceforth "PSNE"), in which, according to the chapter title, "The Supervenience Argument [is] Motivated, Clarified, and Defended". We will attempt to give a reconstruction of the arguments that renders them deductively valid, so that every assumption on which we believe Kim relies is made explicit. We will show that the premises of Kim's arguments are not nearly as innocent as they seem, and that one assumption in particular, which concerns the identity conditions of events, can only be viewed as begging the question against nonreductive physicalism, as that position has traditionally been conceived. Whatever else they may be, Kim's arguments then are not *reductios* of nonreductive physicalism. Nor do they succeed in posing a problem about the possibility of mental causation by showing that we must either accept epiphenomenalism or reject one or another of a set of *prima facie* plausible metaphysical principles. Kim takes himself to have shown that the claims he calls his "substantive premises", which are indeed *prima facie* plausible—at least relative to prevailing assumptions about causation, on which we will comment in the final section of this paper—are not consistent with the claim that mental events have causal efficacy. But what he has *in fact* shown is the inconsistency of a larger set of claims, some of which have no particular *prima facie* plausibility, even relative to the prevailing assumptions. We think the aforementioned assumption about the identity conditions of events is the least plausible among these.

2. The "substantive premises"

Kim thinks his arguments show that the following claims cannot all be true:

Irreducibility

No mental property is identical with a physical property.

Supervenience

All properties strongly supervene² on physical properties.³ In other words: necessarily, for all properties P , all objects x , and all times t : if x has P at t , then, for some physical property P' , x has P' at t and, necessarily, for all y and all times t' , if y has P' at t' , then y has M at t' .

Closure

If a physical event has a cause that occurs at a time t , then it has a physical cause that occurs at t .⁴

Causal Efficacy

Mental events sometimes cause other events.

These assumptions, and Kim's arguments, concern two different kinds of entities: properties and events. Like Kim, we assume that events are concrete *instances*⁵ of properties—not “instances” in the sense of objects that have the properties, but instances in something like the sense of *havings of properties*

² We will abbreviate “strongly supervenes” to “supervenes” in the discussion to follow, as we are not discussing any other varieties of supervenience.

³ In PSNE the principle only says that all *mental* properties strongly supervene on “physical/biological” properties, but Kim, and we, are interested in formulating some minimal *physicalist* commitments, and a physicalist had better think that *all* properties strongly supervene on *physical* properties.

⁴ Curiously, the principle Kim calls “Closure” is not a closure principle. To say that the physical domain is causally closed, in the usual technical sense, would of course be to say that *every cause of a physical event* is also a physical event—not an implausible principle, to our mind. Kim, however, thinks that to assume the physical domain to be causally closed in the literal sense would be to beg the question that is at issue in his argument: it would be “like starting your argument with mind-body causation already ruled out, at least for nonreductivists” (PSNE, 51). We think this is incorrect, for reasons that will become evident in sections 5 and 6, but for now we will join Kim in assuming only the weaker principle that every physical event with a cause has a synchronous physical cause, and calling it “Closure”.

⁵ Kim prefers the term “exemplification”. See Kim (1976).

by particular objects at particular times.⁶ We take the following schema (at least when suitably restricted) to be a platitude: the event that is the having of property P by object x at time t exists if and only if x has P at t .⁷

According to Kim, Irreducibility, Supervenience, and Closure are shared commitments of all nonreductive physicalists. Thus Kim is posing the dilemma: either reject nonreductive physicalism or reject Causal Efficacy, viz., accept epiphenomenalism.

But things aren't quite so simple, as there is a further premise in the argument which is not a characteristically nonreductive physicalist assumption:

Exclusion

No single event can have more than one sufficient cause occurring at any given time, unless it is a genuine case of overdetermination.⁸

Kim says this is a “general metaphysical constraint” (p. 22), so presumably he thinks we should accept it whether we are nonreductive physicalists or not. Whether Exclusion is true or not is not perfectly obvious, and we'll return to this matter in section 7. For now let us simply note that the appeal of the principle is perhaps due to the fact that it *sounds* tautologous: it sounds a lot like the claim that every event has at most one sufficient cause occurring at a given time unless it has more than one sufficient cause occurring at that time.

3. *Other general metaphysical constraints*

Though the assumptions Kim names and calls his “substantive premises” (p. 41) end here, his argument requires several other assumptions that, like

⁶ We shall not consider the possibility that events might be construed as “tropes”. For discussion see papers by Francesco Orilia and others in this volume.

⁷ If P is an n -place relation, then x is an ordered n -tuple of objects. A restriction to the schema in terms of the notion of metaphysical contingency is contemplated in note 27. We find this restriction acceptable but not necessary. Our schema is similar to what Kim (1976, 35) calls the “existence condition”, but not quite the same: see note 27.

⁸ This is Kim's formulation exactly (PSNE, 42). Like Kim, we omit the word “sufficient” in the discussion to follow, but when we use the word “cause”, this should be understood as having an implicit “sufficient” in front of it.

Exclusion, are not clearly nonreductive physicalist commitments. We assume that he regards them, too, as “general metaphysical constraints”. One such assumption is what we will call:

No Overdetermination

For all properties F and G , if F is a supervenience base of G , then no event is causally overdetermined by (events that are instances of) F and G .

Like Exclusion, this principle has a tautologous sound to it. When we say that an event is causally overdetermined by two other events, we mean that it has two causes that are *independently sufficient* for its occurrence. “Independent” here means, surely, at least that it would have been *possible* for each to occur without the other (and that if one had occurred without the other, it would have brought about the same effect). But it is a straight logical consequence of the definition of supervenience that instances of properties and their supervenience bases are not “independent” in this sense, as the instantiation of the base necessitates the instantiation of the supervening property.

Our evidence that Kim makes use of No Overdetermination is in the section titled “Why overdetermination is not an option”, in which Kim answers the question, *Why do instances of mental properties and their supervenience bases not overdetermine their effects?* His answer (p. 48):

The usual notion of overdetermination involves two or more separate and independent causal chains intersecting at a common effect. Because of *Supervenience*, however, this is not the kind of situation we have here. In this sense, this is not a genuine case of overdetermination ...

Since Kim thinks it is “because of supervenience” in the case under consideration, we assume he thinks that supervenience *always* precludes overdetermination, and that is what No Overdetermination says.

In addition, Kim must also assume some other principles relating supervenience to causation. We believe that he is assuming these:

Supervenience-Causation (SC) I

If c causes e and e' is a supervenience base of e , then c causes e' .

Supervenience-Causation (SC) II

If c causes e and c' is a supervenience base of c , then c' causes e .

These principles are perhaps more clearly expressed by means of a visual aid: see Figures 1 and 2. (We represent causation with a single arrow and supervenience with a double arrow pointing from the supervenience base to the supervening event.) SC I and SC II respectively tell us that when the top single arrow in each of Figure 1 and 2 occurs, so does the bottom single arrow.

Figure 1: Supervenience-Causation I

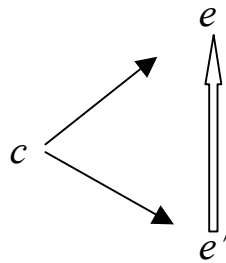
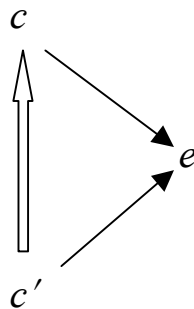


Figure 2: Supervenience-Causation II



One further remark about what SC I and SC II mean is in order. The principles speak of supervenience bases of *events*, though the usual notion of supervenience is defined for *properties*. What we mean by “ e' is a supervenience base of e ”, when e and e' are events, is that e and e' are simultaneous events involving the same object(s) and e' is an instance of *some* property which is a supervenience base of a property that is instantiated by e . More precisely:

The event in which x instantiates P at t is a supervenience base of the event in which y instantiates Q at t' if and only if $x = y$, $t = t'$, and P is a supervenience base of Q .

We believe that Kim introduces SC I at pp. 39-41, where he expresses the principle by saying that an instance of a property can only cause an instance of a supervenient property “by causing its supervenience base” (p. 40).⁹ The grounds for this claim have to do with the alleged fact that there are “two seemingly exclusionary answers” to the question “What is responsible for, and explains, the fact that [e] occurs on this occasion?” These two answers are:

- (a) Because c caused e .
- (b) Because e' , a supervenience base of e , is instantiated on this occasion.

Kim says that a “tension [is] created by (a) and (b)” and that a “simple and natural way of dissipating” this tension is to conclude that c is a cause of e . It is not clear to us why these answers are “seemingly exclusionary”, or even *that* they are seemingly exclusionary (they don’t strike us as seemingly exclusionary),¹⁰ but we will leave it to others to reconstruct and evaluate the reasoning that leads Kim to accept SC I, because an adequate reconstruction of it would require a separate paper¹¹ and we are not, at any rate, singling this assumption out for criticism.

⁹ This does not have exactly the same meaning as our SC I, but we assume that whatever additional import Kim’s “by” has is irrelevant to the arguments; the assumption that the cause of an event e is also the cause of any event subvening e suffices for the purpose for which Kim needs the principle he expresses using the word “by”—namely, deriving step (3) in the two versions of the argument below. Since Kim’s principle differs from our SC I only in being *stronger* than it (“ x does A by B -ing” entails “ x B s”), and the weaker SC I can fill the same role in the argument, we think we can safely ignore Kim’s “by”.

¹⁰ See Marras 2007, section 3, for a discussion of related issues.

¹¹ At the bottom of p. 39 and the top of p. 40 Kim gives a highly impressionistic argument involving counterfactuals and modal operators for the conclusion that c ’s occurrence must have had “something to do with” e' . The argument is open to several different interpretations, which we cannot give here for lack of space. Here the conclusion we are

The reasoning supporting SC II is more transparent. SC II is introduced, obliquely, at p. 41, where Kim states that “There are strong reasons for thinking that [in the kind of situation represented in Figure 1] [c'] is a cause of [e']”. These reasons are:

[c'] is (at least) nomologically sufficient for [c], and the occurrence of [c] on this occasion depends on, and is determined by, the presence of [c'] on this occasion. Since *ex hypothesi* [c] is a cause of [e], [c'] would appear to amply qualify as a cause of [e] as well.

The idea appears to be this: *c supervenes on c' ; therefore c' is at least nomologically sufficient for c* (this depends on reading the innermost modal operator in the definition of supervenience as “at least” nomological necessity). *By assumption, c is a cause of e . Any event that is nomologically sufficient for the occurrence of a cause of an event is also a cause of it; therefore c' is a cause of e' .*

The final “general metaphysical constraint” required by the argument, according to our reconstruction, is:

Closure-Overdetermination

If a physical event p has a cause c that occurs at t , then p has a physical cause p' that occurs at t such that p is not overdetermined by c and p' .

Kim does not express Closure-Overdetermination anywhere in the text, but we attribute it to him because 1) it very neatly fills a gap in one of his arguments, and 2) it is highly plausible *given Closure*, which he does accept. For suppose that a physical event p has a cause c that is *not physical*. Then by Closure p has a physical cause p' synchronous with c . Is p overdetermined by c and p' ? If it is, then *every* event that has a nonphysical cause is overdetermined, and this seems absurd. It is of course logically possible that the physical causes posited by Closure *sometimes* but not always overdetermine their effects jointly with their concurrent nonphysical causes,

invited to draw is that this “something” is the causal relation. The final step in the argument, then, appears to be a kind of inference to the best explanation.

but this seems bizarre and arbitrary. The most natural conclusion is that Closure-Overdetermination is true.¹²

4. *The arguments*

Before giving us his arguments, Kim announces (p. 39):

Properties as such don't enter into causal relations; when we say "M causes M*", that is short for "An instance of M causes an instance of M*" or "An instantiation of M causes M* to instantiate on that occasion." Also for brevity we suppress reference to times.

We will follow Kim in adopting these abbreviating conventions whereby "an instance of property *M*" and the like will be replaced by simply "*M*" and the like, as if properties were identical to their instances. In reading the argument to follow, it will be useful to pretend that the world is such that there are just two times, call them "*t*" and " Δt ", and that anything that is a cause of anything else occurs at *t*, and anything caused occurs at Δt —references to times thus becomes unnecessary.

Kim offers two versions of his argument, and we will discuss them in order. Our exposition differs from Kim's in two respects: First, we present his "versions" of the argument as two distinct arguments, rather than as a "Stage 1" followed by two alternative "completions", as he does. Secondly, we will eliminate certain unnecessary steps, which we will comment on after we give each reconstruction. We will use sorted variables, so that "*P*"s, supplemented with primes as needed, are variables for physical properties and "*M*"s, likewise supplemented, are variables for mental ones.

¹² Compare Papineau (1993, 22-23), who also endorses this principle, effectively giving it the same justification. Papineau correctly observes that Closure-Overdetermination together with the claim that every mental event causes some physical event implies that we must either accept the token (not type) identity of mental and physical events or reject Exclusion (he does not use these labels, but this is what his argument amounts to).

4.1. *Kim's first argument*

In the first argument we introduce for a *reductio* the assumption that a case of mental-to-mental causation occurs:

(1) M causes M' .

By Supervenience, we conclude that:

(2) There is a physical property P' such that P' is a supervenience base of M' .

By SC I we conclude from (1) and (2) that:

(3) M causes P' .¹³

Again, by Supervenience:

(4) There is a physical property P such that P is a supervenience base of M .

By SC II, it follows from (3) and (4) that:

(5) P causes P' .

But Irreducibility tells us that:

(6) $M \neq P$.

From No Overdetermination and (4) we get:

(7) P' is not causally overdetermined by M and P .

¹³ This is not exactly what Kim's step (3) says: it says "M causes M* *by* causing its supervenience base P*" (p. 40, Kim's italics), but see our note 9.

Applying Exclusion to (5), (6), and (7), we conclude that:

(8) M does not cause P' .

(8) contradicts (3), and we have a refutation of (1). In other words, we have derived (9) from our assumptions.

(9) It is not the case that M causes M' .

Since M and M' were arbitrary mental properties, we must conclude the universal closure of (9):

(10) For all mental properties M and M' , it is not the case that M causes M' .

In other words, there is no mental-to-mental causation. To show that there is no mental-to-*physical* causation, we need a separate argument, which can be obtained from the argument just given by deleting lines (1) and (2) and taking (3) (“ M causes P' ”) to be the *reductio* hypothesis. These arguments together commit the nonreductive physicalist to a pervasive epiphenomenalism: mental properties do not cause *any* properties, mental or physical, to be instantiated.¹⁴

The argument just given fits into a larger *reductio*: one against the “substantive premises” of section 1 and the “general metaphysical constraints” of section 2. The conclusion (10) directly contradicts Causal Efficacy. We seem to have a demonstration that the “substantive premises” and the “general metaphysical constraints” cannot *all* be true. Supposing the metaphysical constraints to be off the table, we seem to be forced to reject either Causal Efficacy or one of the other “substantive premises”, which

¹⁴ Or so Kim seems to think. If you think, as we do, that there are properties that are neither mental nor physical (say, biological properties), then you won't agree that a pervasive epiphenomenalism is established yet. However, the argument goes through no matter what property—biological, chemical, aesthetic, whatever—is considered in place of M' .

characterized nonreductive physicalism—in other words we must either accept epiphenomenalism or reject nonreductive physicalism.

The above reconstruction differs from Kim's presentation of the argument in one important respect: in Kim's presentation, Exclusion tells us that "we must eliminate either *M* or *P* as [*P*']s cause", whereafter Closure is called upon to tell us that it must be *M* that gets "eliminated" (pp. 42-43). But as can be seen above, Exclusion together with the preceding three lines yields the conclusion that *M* is not a cause of *P*, contradicting (3) [in Kim's numbering, (8) (p. 41) contradicts (5) (p. 43)]. Citing Closure after a contradiction has been derived cannot contribute anything to the argument. Closure, then, is not actually needed as a premise in the first version of Kim's argument.¹⁵

4.2. *Kim's second argument*

Let us now turn to the second version of Kim's argument. What follows is more paraphrase than reconstruction. Our lines (1) and (3)-(6) come from Kim more or less *verbatim* except for a change in the numbering, and as noted in note 7 above; the remaining lines are close paraphrases of lines in Kim's argument. Lines (1)-(3) are what Kim calls "Stage 1", and lines (4)-(7) he calls "Completion 2". These taken together comprise Kim's second argument. As before, we make for a *reductio* the assumption that there is a case of mental-to-mental causation:

(1) *M* causes *M'*.

¹⁵ It might be suggested that, instead of SC II and No Overdetermination, Kim is tacitly appealing to Closure-Overdetermination, as we believe he is in his second argument (see below), to derive steps (5) and (7). However, there is no textual evidence for this. At pp. 41-42, where these steps are derived, Kim makes no mention of a causal closure principle (except parenthetically at the top of p. 42, to announce that Closure will *later* be applied to "disqualify *M* as a cause of [*P*']"—the step that according to us is unnecessary). What Kim *does* cite here is the fact that *P* is "at least nomologically sufficient" for *M* (quoted above), and it is clear that he thinks this is so because *M* supervenes on *P*. This, and Kim's later remarks about supervenience precluding overdetermination (p. 48), is to us decisive evidence that SC II and No Overdetermination are the operative assumptions here.

Again we know by Supervenience that:

- (2) There is a P' such that P' is a supervenience base of M' .

From (1) and (2) it follows by SC I that:

- (1) M causes P' .

By Closure, (3) implies that:

- (4) P' has a physical cause—call it P —occurring at the time M occurs.

But Irreducibility tells us that:

- (5) $P \neq M$.

Next Kim asserts a line without justification [he introduces it only with “hence” (p. 44)]:

- (6) P' is not overdetermined by P and M .

Now, from (4), (5), (6) and Exclusion it follows that:

- (7) M does not cause P' .

(7) contradicts (3), and once again we must conclude that M did not cause M' . As M and M' were arbitrary mental properties, the conclusion must be that no mental-to-mental causation occurs ever, and the conclusion that no mental-to-physical causation occurs either follows as in the second variant of the first argument.

In addition to the large problem shared by both versions of Kim’s argument discussed in the next two sections, there are two minor problems specific to the second version, as presented by Kim.

The first problem is that not only is line (6) not justified, but it *can't* be justified from assumptions explicitly made by Kim, his “substantive premises”. This is why we attributed Closure-Overdetermination to Kim. It would enable him to complete the argument as follows.¹⁶

⋮

(4) P' has a physical cause—call it P —occurring at the time M occurs—and P' is not overdetermined by P and M . [From (3) by Closure-Overdetermination.]

(5) $P \neq M$. [From Irreducibility.]

(7) M does not cause P' . [From (4) and (5) by Exclusion.]

The second problem is not so easy to solve on Kim's behalf. It is that Kim says that in his second argument “Supervenience is not needed as a premise” (p. 44), but this is not so: the second argument shares what Kim calls “Stage 1” with the first argument, and Supervenience is needed in Stage 1. Without Supervenience, we cannot conclude that M' has a physical supervenience base P' . According to our reconstruction, the difference between the two

¹⁶ There are, to be sure, other ways to complete the argument on Kim's behalf, which are not implausible. The alternatives that come to mind, however, would have Kim appealing to the supervenience of M on P , whereas he insists that his second “completion is simpler than Completion 1” in part because “Supervenience is not needed as a premise” here (p. 44). The main alternative we can think of for Closure-Overdetermination is:

Closure-Supervenience

If a physical event p has a cause c occurring at a time t , then p has a physical cause p' such that 1) p' occurs at t and 2) p' is a supervenience base of c .

This principle could be given a justification similar to the one we gave Closure-Overdetermination: it could be argued that if the principle is not true, absurdly pervasive overdetermination will follow (unless mental events happen to be identical to physical events—a possibility we will consider later). For lack of space, we leave working out the exact justification for Closure-Supervenience, as well as how Kim's second argument could be completed using it, as an exercise for the reader.

arguments is different from what Kim asserts: they differ with respect to two premises: the first argument makes use of SC II and No Overdetermination whereas the second does not; the second argument makes use of Closure-Overdetermination whereas the first one does not. It is, of course, entirely possible that our reconstruction does not capture Kim's intentions (see note 15), but our goal has been to produce a reconstruction that is as close to the letter of Kim's exposition as possible.

5. *The main logical problem*

Here is the most serious problem with Kim's arguments. That the arguments appear to have a valid form is only an artifact of the abbreviating convention used by Kim—unabbreviated, the arguments are either invalid or enthymematic; we suppose the latter. We followed Kim in using the same variables for both properties and events, but now we will adopt a more explicit nomenclature, with uppercase letters ("*M*" and "*P*", possibly supplemented with primes) for properties and lowercase letters for their instances ("*m*" and "*p*", possibly supplemented with primes). Using this convention and adding material presumably elided by Kim, the steps of our first reconstruction are as follows, where the added material is in **boldface**.

- (1) **An instance *m* of *M* causes an instance *m'* of *M'*** [Assumption].
- (2) There is **an instance *p'* of a physical property *P'*** such that *p'* is a supervenience base of *m'*. [From (1) by Supervenience.]
- (3) *m* causes *p'*. [From (1) and (2) by SC I.]
- (4) There is **an instance *p* of a physical property *P*** such that *p* is a supervenience base of *m*. [From (1) by Supervenience.]
- (5) *p* causes *p'*. [From (3) and (4) by SC II.]
- (6) *M* ≠ *P*. [By Irreducibility.]
- (7) *p'* is not causally overdetermined by *m* and *p*. [From (4) by No Overdetermination.]
- (8) *m* does not cause *p'*. [From (5), (6), and (7) by Exclusion.]

It is clear that (8) cannot be derived from (5), (6), and (7) by Exclusion: for (8) to be so derivable, (6) would have to say “ $m \neq p'$ ”, not “ $M \neq P'$ ”. Nor does (8) follow from previous steps by any of the other assumptions that we have made explicit so far. However, one interesting conclusion does follow, namely:

(8') $m = p$.

The argument is straightforward: since m and p are simultaneous causes of p' [by (3) and (4)], and m and p do not jointly overdetermine p' [by (7)], it follows by Exclusion that $m = p$.

Since we assumed nothing about m except that it is a cause of some other mental event, what we have here is a proof that any mental event that causes another mental event is identical to a physical event. If we assume, instead of (1) that m causes a *physical* event, we can show that any mental event that causes a physical event must also be identical to a physical event. So we have shown that any mental event that causes any other event—mental or physical¹⁷—must be identical to a physical event. We have, in other words, a new argument for the *token identity theory* (or *token physicalism*) first proposed in Davidson 1970. Unlike Davidson's argument for the theory, however, this one does not assume what Davidson called the “nomological character of causality”—that causes and effects must be related by deterministic laws—which many philosophers now find implausible, so perhaps this argument is more compelling than Davidson's.

But clearly Kim does not think he is giving us an argument for the token identity theory. He thinks he is showing the inconsistency of Causal Efficacy with the assumptions he attributes to the nonreductive physicalist and certain “general metaphysical constraints”. So what has gone wrong? It would be most implausible to suggest that Kim has simply confused properties with their instances because of his abbreviating convention. Rather, we suggest that Kim is making use of an unstated premise which he likely thinks of as yet another “general metaphysical constraint”. It's not difficult to see what this hidden premise might be, as Kim makes it explicit in his earlier work on

¹⁷ Again, this dichotomy may not be exhaustive, but the argument goes through no matter what kind of property we consider in place of M .

events. It is an assumption about the identity conditions of events. To express this assumption we must first introduce a bit of notation from Kim 1976: let

$[x, F, t]$

denote the event (if any) in which object x possesses property F at time t .¹⁸ Using this notation, Kim (1976, 35) states what he calls the “identity condition”:

Identity Condition

$[x, P, t] = [y, Q, t']$ just in case $x = y$, $P = Q$, and $t = t'$.

The Identity Condition implies that if F and G are distinct properties and f and g are instances of F and G respectively, then $f \neq g$. Given this, we can conclude from steps (1), (4), and (6) of the first argument that $m \neq p$, and we can complete the argument:

∴

(7) $m \neq p$. [From (1), (4), and (6) by the Identity Condition.]

(8) p' is not causally overdetermined by m and p . [From (4) by No Overdetermination.]

(9) m does not cause m' . [From (5), (7), and (8) by Exclusion.]

With the introduction of the Identity Condition, however, the logical problem has been transformed into a dialectical problem for Kim.

6. *The dialectical problem*

The dialectical problem is that if Kim does include the Identity Condition among his assumptions, he can no longer claim to have shown that Causal Efficacy is inconsistent with *nonreductive physicalism* plus his metaphysical

¹⁸ Again, if F is an n -place relation, x is an ordered n -tuple of objects.

constraints. The Identity Condition implies that each event is an instance of exactly one property¹⁹—which is expressed using Kim’s notation as:

(F) If $[x, F, t] = [y, G, t]$, then $F = G$.

Let us call any view of events that is committed to (F) *fine-grained*, and let us call any view of events committed to (F)’s denial *coarse-grained*. Let us now consider the question: *What is nonreductive physicalism?* According to one popular answer, nonreductive physicalism just is the conjunction of token physicalism with the denial of type physicalism.²⁰ But to have a fine-grained conception of events is to assume that token physicalism implies type physicalism,²¹ which is—if you accept that popular answer—well nigh synonymous with the statement that *nonreductive physicalism is false*. On this conception of what nonreductive physicalism is, Kim’s supervenience argument, if it is to be viewed as an argument against nonreductive physicalism, reduces to the claim that the nonreductive physicalist gets the identity criteria of events wrong, and the various assumptions about causation and supervenience do no work at all.

This is not, of course, the only possible characterization of nonreductive physicalism,²² but it is certainly the *classic* one. Recall that token physicalism was an essential component of the positions defended in the two *ur-documents* of nonreductive physicalism—Davidson 1970 and Fodor 1974. Psychophysical supervenience played no role in their arguments, and it only became a physicalist staple due to Kim’s subsequent work. And although Kim regards psychophysical supervenience as a minimal requirement for any form

¹⁹ Namely, the property that Kim calls the event’s “constitutive” property (as explained in section 7 below).

²⁰ Perhaps more accurately, the conjunction of token physicalism with Irreducibility, which is stronger than the denial of type physicalism. The falsity of type physicalism would only require the existence of *one* property that is not a physical property.

²¹ As Francesco Orilia also notes in his contribution to this volume (section 6). He also points out that Kim (1996, 60) explicitly recognizes this implication of his account of events.

²² There are others that appeal to the idea of “realisation” or “constitution”, and that are compatible with a fine-grained conception of events. Cf. Boyd (1980, 82-87; 101-103), Cummins (1983, 22-23); Papineau (1993, ch. 1).

of physicalism (Kim 1998, 15),²³ at the same time he no longer regards it as a sufficiently robust relation on which to base a physicalist theory of mind, particularly since supervenience can be defined over “multiple domains” (Kim 1988), rendering it compatible with various forms of parallelism (including epiphenomenalism). Notice, however, that if psychophysical supervenience is defined over a *single* domain of *events*, it entails token physicalism, for every mental event will *be* a physical event. At any rate, token physicalism remains a highly plausible candidate for a “minimal” physicalist commitment, and it still enjoys wide acceptance as a position that purposively distinguishes itself from type physicalism. Consequently, to construct an argument against nonreductive physicalism, one of whose premises effectively says that token physicalism implies type physicalism, is to adopt a rather odd dialectical strategy.

We will consider in a moment the likely response that classic (token physicalist) versions of nonreductive physicalism lack the resources to account for mental causation (at least to the extent that they are silent about the causal role of properties), and so cannot be taken seriously as versions of nonreductive physicalism. But first there is a further problem with Kim’s argument that needs to be considered.

7. *A further problem*

The further problem is that the Identity Condition does not appear to be consistent with Exclusion—at least not if we understand “overdetermination” in a particular way, which seems to us natural. Exclusion implies that if an event has two distinct synchronous causes, then it is overdetermined by them. Brutus’s killing of Caesar (call this event “BKC”) and Brutus’s murdering Caesar (call this “BMC”) are synchronous events that have many common effects. But clearly BKC and BMC cannot overdetermine their common effects if we assume, as we have done, that overdeterminers must be

²³ Unsurprisingly, since on his fine-grained conception of events, physicalism must minimally require that mental events, albeit distinct from physical events, somehow “depend on” and be “determined by”, physical events. And “realisationist” versions of nonreductive physicalism obviously imply supervenience, since the realisation relation implies the supervenience relation (as Kim recognizes in his 1998, 23-24).

independent at least in that each could have occurred without the other (the murder could not have occurred without the killing). Now, because murdering and killing are two distinct properties (relations), the Identity Condition implies that $BKC \neq BMC$. But because BKC and BMC are synchronous events with at least one common effect, and BKC and BMC do not overdetermine any of their effects, Exclusion implies that $BKC = BMC$.²⁴ This is not an outright *logical* inconsistency: no contradiction follows from the set {Exclusion, Identity Condition} alone. To avoid having to choose between Exclusion and the Identity Condition, and thus giving up his supervenience argument, Kim can argue that: (a) BKC and BMC are not synchronous, (b) BKC and BMC have no common effects, (c) BKC and BMC do, after all, overdetermine all of their common effects, or (d) the property of killing *is* the same as the property of murdering. None of these seems very promising to us.

There is, however, a fifth alternative: (e) Kim could reply that the event description notation “[x, P, t]” used in the formulation of the Identity Condition is not to be understood in terms of the commonsense idea of an object having a property at a time, but in terms of a technical notion of a “constitutive property” of an event. The idea would be that the “ P ” in “[x, P, t]” always denotes a special, unique property “constitutive” of the event denoted by “[x, P, t]”, and that not every property x has at t can be a constitutive property of an event. It should be clear how this distinction would enable Kim to dodge the objection about BKC and BMC: the application of the Identity Condition in our argument for the distinctness of BKC and BMC would simply be invalid, because the argument does not have premises stating that *killing* and *murdering* are constitutive properties and hence we are not entitled to conclude that there exist any such events as [\langle Brutus, Caesar $\rangle, kills, t$] and [\langle Brutus, Caesar $\rangle, murders, t$] for us to apply the Identity Condition to. (Of course, if the Identity Condition and Exclusion are both to be maintained, it had also better be the case that at least one of *killing* and *murdering* is *not* a constitutive property, or else the argument could be supplemented with additional true premises to yield a refutation of the Exclusion/Identity Condition combination. We do not, however, have any arguments to show that *murdering* and *killing* are constitutive properties.)

²⁴ As one would expect under a coarse-grained conception of events.

We believe that Kim would choose alternative (e) for coping with the present problem—the distinction between constitutive and other properties of events is, in fact, drawn by Kim in his 1976. Although in that paper Kim resists the identification of such events as BKC and BMC,²⁵ he is open to the identification of other events involving the instantiation of distinct properties by an object.²⁶ But choosing alternative (e) comes at a price: To begin, Kim would have to reject the principle that $[x, P, t]$ exists if and only if x has P at t (an unrestricted form of what Kim 1976, 35, calls the “existence condition” for events).²⁷ But if this principle is rejected, the meaning of Kim’s event

²⁵ On p. 44 Kim indicates that he would not identify Brutus’s stabbing Caesar with Brutus’s killing Caesar, and on p. 43 he says that “no stabbings are killings and no killings are assassinations”. These are similar enough to our examples.

²⁶ On p. 44 Kim concedes that Brutus’s stabbing Caesar and Brutus’s stabbing Caesar with a knife *are* the same event and he accordingly denies that the relation x *stabs* y *with a knife* is a constitutive property of the event denoted by “Brutus’s stabbing Caesar with a knife”, though it clearly is a property of the pair \langle Brutus, Caesar \rangle .

²⁷ Kim intends his “existence condition” schema to be understood with the restriction that only predicates that ascribe constitutive properties may be substituted for “ P ” (Kim 1976, 34-37). Our own attitude about what events exist is liberal: we do not think our existence schema “The event in which x has P at t exists iff x has P at t ” (section 2) is in need of restricting. Events in which properties of every kind are instantiated are required for the semantic analysis of natural language. Only some events can enter into the causal relation—only the metaphysically contingent ones, we suppose, viz. events e such that it is metaphysically possible both for e to have occurred and for e to not have occurred. For discussions of causation, our existence schema *could* be restricted so that only predicates that ascribe contingent properties (properties such that possession of them is a metaphysically contingent matter) may be substituted for “ P ”, and the semanticist’s events could be called something other than “events” (“schmevents”, for example), but this is a question of word choice, not substance. Such a restriction is intelligible inasmuch as the notion of metaphysical possibility is, and, it seems to us, it would correctly disqualify as events those schmevents that cannot enter into the causal relation.

Note that the question whether to restrict the schema

(E1) The event in which x has P at t exists iff x has P at t ,

which we used in section 2, is different from that of whether to restrict the schema

(E2) $[x, P, t]$ exists iff x has P at t ,

description notation is no longer clear: to understand it we would need an account of just what it is that makes some properties “constitutive” and others not, and none has been provided by Kim. We do not claim that such an account cannot be provided—the salient point is, rather, that for anyone who is worried about the coherence of the Exclusion/Identity Condition combination, as we are, the acceptability of Kim’s supervenience arguments will depend on his ability to produce such an account.

What’s more, if Kim chooses (e), a new logical problem is introduced into the argument. The problem is that option (e) renders not only *our* argument against the coherence of the Exclusion/Identity condition combination *but also Kim’s own supervenience arguments* invalid, for exactly the same reason: if the Identity Condition is to be applied to m , M , p , and P in the supervenience arguments, the arguments will require two additional premises to remain valid: 1) that M is a constitutive property of m , and 2) that P is a constitutive property of p . Perhaps it will turn out that both claims fall out of the correct account of constitutive properties—even though since writing his 1976, Kim has indicated that, on pain of having to revise his property-exemplification account of events, he may have to deny that mental properties can be constitutive properties of events!²⁸ However that may be, it

which is what Kim is concerned with in his 1976. The latter is a partial definition of a new piece of notation (“[”, “]”), call this the *canonical event description* notation. Ruling out a particular predicate “ R ” as an allowable substituent for “ P ” in (E2) would *not* have the consequence that there is no such thing as the event in which x has R at t ; it would only have the consequence that that event, if it exists, is not denoted by “[x , R , t]”—it might be denoted by some *other* canonical event description. Presumably Kim would not impose the same restriction on the ordinary language (E1) as on the quasi-formal (E2), since in 1976 he wants to allow events to instantiate properties which are not constitutive of them (see our note 24). The restriction Kim proposes for (E2) differs from the one just contemplated for (E1) in that the former is not made using terms of which we have a prior understanding but by using the technical term “constitutive property”, which Kim does not explicitly define (see pp. 36-37).

²⁸ Kim notes that “a revision of the standard property-exemplification account of events (essay 3) [Kim 1976]” may be called for, “especially if mental properties, in spite of their multiple realisability, are accepted as legitimate event-generating properties. For on the standard account two property instances count as distinct events if the properties instantiated are distinct *Considerations advanced in Kim (1992) concerning disjunctive properties may be reason enough for excluding mental properties as constitutive properties of events*” (Kim 1993, 364-65, note 5, our italics).

should be clear that if option (e) is chosen, some further work on Kim's part will be required to make the supervenience arguments convincing.

8. *A problem about mental quausation?*

Kim is well aware, of course, that there are token physicalists who are not type physicalists. Does he think his supervenience argument has anything to say to them? Surprisingly, he does. In footnote 9 in ch. 2 of PSNE and elsewhere (e.g. Kim 1998, ch. 4) Kim hints that his argument could be reconstructed so as to have bite against nonreductive *token* physicalist positions. Unfortunately, however, Kim never gives us an explicit reconstruction. Though it would be obviously unfair to criticize an argument never explicitly formulated, we would still like to register the reason for our scepticism that a plausible argument along the lines Kim hints at can be made. Kim says (PSNE, 42, n. 9) that in the reconstructed argument,

“An *M*-instance causes a *P*-instance” must be understood with the proviso “in virtue of the former being an instance of *M* and the latter an instance of *P*”.

The argument would then be one about *quausation* (to use Terence Horgan's (1989) term): causation *qua* something.²⁹ But then we would need a *quausal exclusion* principle to replace Exclusion. What might it be? The trouble is that Kim's suggestion does not determine a unique translation of Exclusion into quausal terms, and we are left wondering what principle he might have had in mind. The most *straightforward* rendering of Exclusion into quausal terms that we can think of is:

²⁹ The closest Kim comes to providing an argument is in the following: “Suppose that a certain event, in virtue of its mental property, causes a physical event. The causal closure of the physical domain says that this event must also have a physical cause. We may assume that this physical cause, in virtue of its physical property, causes the physical event. The following question arises: *What is the relationship between these two causes...?*” (1989, 280). Kim goes on to suggest that, barring overdetermination and given closure, the physical cause excludes the mental one unless the mental and the physical *properties* of the two causes are identified. The implication here is that the mere “token identity” of two causes as coarse-grained events won't give us a solution to the problem; what needs to be identified is “that in virtue of which” the cause causes what it does.

Quausal Exclusion 1

Except in cases of overdetermination, if F and G are distinct properties, and an instance c of F causes an event e in virtue of c 's being an instance of F , then it is not the case that c causes e in virtue of c 's being an instance of G .³⁰

This principle, however, has no intuitive support that we can find, and we think it would be unfair to attribute it to Kim. Similar principles have been ably criticized by Fodor (1989), Jackson and Pettit (1990; see (B) at p. 110 for a straightforward counterexample), Heil and Mele (1991), and Yablo (1992), and we have little to add to those critiques beyond pointing out that a problem analogous to the one considered earlier about BKC and BMC would arise again: given Quausal Exclusion 1, where e (for example, Calpurnia's grieving) is an effect of Brutus' action of killing/murdering Caesar, BKC causing e in virtue of BKC's being a murdering of Caesar would exclude BKC's causing e in virtue of BKC's being a killing of Caesar—surely something we should have no reason to claim.

Let us instead look to Kim's other writings to see if a quausal exclusion principle might be *derived* from other claims he has made. In Kim 1988 we find a "principle of explanatory exclusion" (PEX), which states: "No event can be given more than one *complete* and *independent* explanation" (p. 239). There Kim also concedes that PEX is "something that many will, I'm afraid, consider absurdly strong and unacceptable" (*ibid.*). Absurdly strong or not, if we also assume Kim's "explanatory realism",³¹ we could try to make a case for the following principle on the basis of PEX:

Quausal Exclusion 2

Except in cases of overdetermination, if F and G are distinct *and independent* properties, and an event c causes another event e in virtue of c 's being an

³⁰ This is essentially the principle attributed to Kim by Marras, 2000, p 145, as principle (Q*). Note that in addition to a quausal exclusion principle, the reconstructed argument would also require analogous revisions of Closure and/or of SC I and SC II, as well as of any premise to the effect that a mental (physical) event causes another event. (E.g., Closure would become Quausal Closure: "If a physical event e has a cause that occurs at time t , then it has a physical cause c that occurs at t and that causes e in virtue of some physical property of which c is an instance".)

³¹ Which is itself far from obvious. See Marras 1998 for discussion.

instance of F , then it is not the case that c causes e in virtue c 's of being an instance of G .

For lack of space we will not attempt to show just *how* Quausal Exclusion 2 might be derived from PEX; the interested reader may consult the argument of Macdonald and Macdonald (2006, 544-545), which uses PEX to justify a principle very similar to Quausal Exclusion 2. Rather, our main point about Quausal Exclusion 2 is that, *however plausible or implausible PEX may be, and whatever difficulties might be involved in using PEX to justify Quausal Exclusion 2, if Quausal Exclusion 2 is the principle that will be used in the new supervenience argument, then there is a straightforward reply*. The reply has been around in the mental causation literature, in various forms, for a few years.³² It is simply this: Quausal Exclusion 2 cannot be used to argue that the causal efficacy of mental properties is “excluded” by that of their supervenience bases, because such an application of Quausal Exclusion 2 would require mental properties and their supervenience bases to be *independent*. By the very definition of supervenience, there is a *necessary connection* between mental properties and their supervenience bases, so mental properties are not independent of their supervenience bases.

The “quausal” formulation of the mental causation problem aims to highlight the point that mental causation “ultimately involves the causal efficacy of mental *properties*” (Kim 1998, 37), and Kim’s challenge to the token physicalist is to explain how a mental event, even if token-identical with a physical event, can cause what it does in virtue of the mental properties it instantiates (where these are distinct from the physical properties it instantiates), and why the causal efficacy of the mental properties is not preempted by the efficacy of the physical properties. However, to merely complain that token physicalism, as such, lacks the resources to account for the efficacy of mental properties³³ is not to the point: the relevant issue is

³² See Macdonald and Macdonald 2006, 566, Bennett 2003, and Papineau 1993, ch. 1, sections 6 and 7.

³³ This was, essentially, Kim’s (1993) complaint against Davidson’s anomalous monism. While granting that Davidson was “arguably right” in denying that anomalous monism *entails* the causal inertness of mental properties, he nonetheless insists that anomalous monism “fails to provide mental properties with a causal role” (p. 20). As argued in Marras (1997), this was merely a failure of omission: Davidson’s monism was not intended to provide a theory of the causal efficacy of mental *properties*.

whether token physicalism can *coherently be supplemented* with additional “substantive premises” (e.g., about the physical realisation of higher-level properties, or their implementation in physical mechanisms, etc.)³⁴ so as to give mental properties a causal role—a role which token physicalism as such does not *preclude*.

9. *Is there a problem about token identity?*

Why has the fact that Kim’s supervenience argument begs the question against nonreductive physicalist positions that assume token identity not received wide attention? There are two fairly common beliefs that might be responsible for this. One is that “the problem of how events should be individuated” “plays no essential role in questions about epiphenomenalism” (Shapiro and Sober forthcoming, n. 3)—and this is just what Kim appears to think. The other is that there is something wrong with token physicalism, and that nonreductive physicalism ought to be formulated in terms of a Kim-style fine-grained conception of events. Our response to the first belief is implicit in what we said in the previous two sections: the supervenience arguments Kim *actually* gives us depend essentially on an assumption he makes about the identity criteria of events, and it is questionable whether Kim can construct a convincing alternative to the actual supervenience arguments without making that assumption. Clearly, then, the way we individuate events does play an essential role at least in the questions about epiphenomenalism raised by *Kim*. Our response to the second belief is that there quite simply are no good arguments against token physicalism; or, less tendentiously and more to the point, that there are no good arguments against token physicalism that are not equally good arguments against type physicalism.³⁵ (This is

³⁴ See, e.g., Fodor 1989, McLaughlin 1989, Jackson and Pettit 1990, Hardcastle 1998.

³⁵ Kripke’s well-known modal arguments against the identity theory, for example, can equally be directed against the type and the token identity theory. So can any of the familiar arguments based on the conceivability/possibility of zombies and/or disembodied minds. Such arguments must of course be distinguished from arguments against *nonreductive* (versus *reductive*) physicalism, either of the token identity variety or of the realisation variety. These latter arguments need to show *not* that token identity is false, but that the *denial of type identity* is false. Typically these aim to show that the multiple

unsurprising, since type physicalism *entails* token physicalism; so if the latter is shown to be false, so is the former.) The arguments against token physicalism are notoriously controversial,³⁶ and at any rate, they are of no help to a friend of *type* physicalism like Kim since, as just remarked, if any of these arguments defeat token physicalism, they defeat type physicalism.³⁷

realisability thesis is false, or that it does not stand in the way of a reductive account of mental properties. See, e.g., Kim 1992, Richardson 1979, Bickle 1998.

³⁶ For a recent, forceful critique of such arguments see Papineau 2002.

³⁷ Consider, for example Burge's (1979) argument against token physicalism. The argument is something like this: Tokens of those mental states that are broadly individuated cannot be identical to tokens of the physical states that realize them because the realizing properties are plausibly *intrinsic* ones. Suppose, for example, that Joe's belief that *p* at time *t* is realized by Joe's brain state *S* ("Joe's *S*" for short). If the content of Joe's belief is broadly individuated, then there is a possible world *w* in which Joe is intrinsically exactly like he is in the actual world, so he has *S* at *t* in *w*, but in which Joe's does not believe that *p* at *t* because of some difference in his environment. Now it would seem that on the basis of this we can argue:

- (1) It is possible Joe's belief that *p* occurs but Joe's *S* does not occur (premise).
- (2) It is possible that Joe's belief that *p* ≠ Joe's *S* (from (1) by Leibniz's law and propositional modal logic).
- (3) Joe's belief that *p* ≠ Joe's *S* (from (2) because objects that are possibly distinct are actually distinct).

The argument can clearly be reformulated so as to apply not just to Joe's (token) belief that *p* and to Joe's *S*, but also to the types (properties) *believing that p* and *being in S*. After all, "It is possible that Joe's belief that *p* occurs but..." is necessarily equivalent to "It is possible that *believing that p* is exemplified by Joe but..."

That having been said, we note that the argument is *not* particularly convincing when applied (as intended) against *token* physicalism. That is, while it is true that *if* the argument is successful in refuting token physicalism, it is also successful in refuting type physicalism, we do not think that the argument *is* successful against token physicalism. The trouble is that, as a matter of modal logic, in order for the move from (1) and (2) to (3) to be valid, or, equivalently, in order for ' $\Diamond a \neq b \rightarrow a \neq b$ ' to be a logical truth, '*a*' and '*b*' must be *rigid designators*. However, the token events at issue in the argument are denoted by *definite descriptions*, which are *not* rigid—at least not rigid *de jure*. Definite descriptions may, however, be rigid *de facto* if they happen to pick out their designata by their essential properties (e.g. 'the even prime' is rigid *de facto*). Burge of course recognizes this—he claims that tokens of beliefs have their contents essentially (Burge

The reasons that have led some philosophers to prefer “token realisationism” to token identity physicalism are far from persuasive, and/or are largely motivated by a prior commitment to a fine-grained conception of events. Boyd (1980), for example, claims that his realisationist version of nonreductive physicalism enables him to cope with Kripke’s arguments against identity theories. But he himself shows, in the same paper, how identity theories are quite able to resist Kripke’s arguments; so what makes realisationism preferable to token identity? And Papineau (1993, 24) explicitly acknowledges that his reason for preferring realisationism is that, *qua* terms of causal relations, events are best viewed as fact-like entities, or as structured, fine-grained events à la Kim. If so, how we individuate events again does, *pace* Kim, play a role in the mental causation debate.

1979, 75), but this is far from obvious. In fact, the very externalist considerations Burge brings to bear on token physicalism seem to militate against this conclusion. It seems to us entirely natural to describe the moral of Burge (1979) by saying things like: “There are counterfactual circumstances in which Joe’s belief that *p* at *t* would have had a different content”, “If the stuff in the oceans had been XYZ, then (the present token of) my belief that water is wet would have had a different content than it in fact does”. (Note that the argument would be valid if the descriptions in (1) were read as having wide scope relative to the modal operator. However, if (1) were read this way, it would receive no support from Burge’s externalist considerations: the latter only show that the mental properties of a person can come apart from the physical properties that, in the actual world, realize them. They do not show that the *events* in which these properties are instantiated in the actual world are *distinct* in some other world.)

If, on the other hand, the argument is made using definite descriptions that designate *properties* instead of events, there is no problem about rigidity. Any property—at least any property we have a predicate for—can be designated by a rigid definite description that makes use of the predicate we would normally use to ascribe the property. For example “the property of being red”, or “the property *red*” are rigid: in no world does “the property of being red” designate any property other than redness. Similarly, “the property of believing that *p*” in no world designates anything other than the property of believing that *p*.

10. *Concluding (polemical) remarks*

Our critique of Kim's exclusion argument aimed to show that the reasons he gives for holding that nonreductive physicalism entails epiphenomenalism are not convincing, resting as they do on a conception of events that rules out from the start a classic and perfectly coherent version of nonreductive physicalism that we believe still merits consideration—a conception of events that is, furthermore, only dubiously consistent with another central premise of his arguments. To have accomplished this aim is not, of course, to have provided a positive account of how mental causation is possible within the confines of nonreductive physicalism thus construed. To provide such an account one would have to explain, minimally, how mental events can have causal efficacy not merely in virtue of their being token-identical with physical events, but also in virtue of the mental properties they instantiate. Although the question of how mental causation is possible is surely a legitimate one, it is not one that we need to address here; it suffices to have shown that there is no compelling argument that Kim (or anyone else, to our knowledge) has provided for the conclusion that a satisfactory account of mental causation *cannot* be provided within the bounds of nonreductive physicalism.

Still, while the question of how mental causation is possible is legitimate and important, there is something perplexing about the almost exclusive attention that it has received from Kim and others for so many years. There is no *prima facie* reason to suppose that the question of how *mental* causation is possible is more special, or more difficult, than the question of how *biological*, or *chemical*, or, indeed, *physical* causation is possible.³⁸ Moreover, Kim's worry that physical causation may somehow "exclude" mental (or other higher-level) causation presupposes that there is *physical* causation, and that it is somehow less problematic than mental or other higher-level causation. These presuppositions, however, cannot be taken for granted. In fact, we think this stance gets matters backwards: if the reality of any causation can be taken for granted, we think the reality of causation at

³⁸ That Kim thinks that the problem of mental causation poses special difficulties is evident in ch. 3 of his 1998 book, especially in view of his claim that the supervenience argument "does not generalize" to other domains beyond the mental.

higher levels, including the mental, should be, whereas the reality of physical causation should be considered an open question.³⁹

The reason is this. The concept of causation we are interested in is not a philosophers' invention; it does not get its content from metaphysical principles—Closure and others—of the sort that are discussed in the literature on mental causation, but from its use in *science* and *common-sense understanding*. And if we are interested in finding out what really causes what, we should listen to what science in particular has to say about this matter. Now it happens that the concept of cause is found *only* in the special sciences, and *never* in fundamental physics. That being so, we should accept the reality of higher-level causation (the kind that the special sciences talk about), including mental causation, and remain noncommittal, until some arguments are provided, on the reality of physical causation. Unless physics changes dramatically and its practitioners begin to talk about causation, these arguments will have to be *philosophical* arguments. Perhaps the most straightforward argument for physical causation is one that assumes token physicalism and the reality of causation at some higher level, call it *L* (*L* could be, but doesn't have to be, the mental level): all *L*-causes are events, and all events are physical, so some physical events are causes. But this still leaves open the question of whether physical events cause other events *in virtue of* being instances of particular physical properties.

These may seem like strange things to say, but consider the fact that physical events are instances of *physical properties*. We assume that physical properties are those properties that are ascribed by predicates that occur in quantum physics (or in whatever the correct lowest-level theory turns out to be), or by open sentences containing only such predicates. Instances of physical properties are *virtually never* cited as causes, at least not under their physical descriptions. Thus it takes an argument to show that instances of physical properties can be causes. If we assume that all instances of higher-level properties, some of which *are* commonly cited as causes both in (higher-level) science and common sense, are identical to instances of physical properties, then it is clear that there is physical causation all around us; but this still leaves all the physical events that are not identical to instances of any higher-level properties unaccounted for (presumably there

³⁹ For a similar stance and a thorough discussion of the issues, see Ladyman and Ross 2007, ch. 5.

are many such physical events), as well as the causal powers of the physical *properties* themselves. Trying to identify the causes and effects, and the causally efficacious properties, in the world *as described by physics* is not an easy task; it would require us to have a correct and informative analysis of the conditions under which one event can be said to cause another, as well as of the conditions under which a property instantiated in a cause can be said to be causally efficacious, and it's not obvious that we have such an analysis.⁴⁰

Not only is the reality of physical causation less obvious than the reality of higher-level causation, but arguments similar to those found in the mental causation debate could be used for “excluding” physical causation. The problem could be put in terms of a Malcolm 1965-style explanatory exclusion principle: given that we can have, in principle, a complete covering-law explanation for every physical event in the vocabulary of quantum physics, and that these explanations make no use of the concept of cause, what role is there left for physical causation to play? Physical causation is “excluded” by the initial conditions and the laws of physics, which determine, without positing any metaphysical glue binding one event to another, the occurrence (or objective probability of the occurrence) of every physical event.

Not that we think this is a good argument, but its similarity to some of the arguments used to motivate epiphenomenalist worries about mental events and properties should raise some questions about the seriousness of those worries.

REFERENCES

- Bennett, K. (2003). Why the Exclusion Problem Seems Intractable, and How, Just Maybe, to Tract It. *Noûs*, 37, 471-497.
- Bickle, J. (1998). *Psychoneural Reduction: The New Wave*. Cambridge, MA: MIT Press.
- Boyd, R. (1980). Materialism without Reductionism. *Readings in the Philosophy of Psychology, Volume I*, ed. N. Block. Cambridge, MA:

⁴⁰ However, there are some promising developments: recent attempts to explicate causation in terms of the outcomes of counterfactual interventions (Woodward 1997, Pearl 2000) might provide us with the means to do just this.

Harvard University Press, 67-106.

Burge, T. (1979). Individualism and the Mental. *Midwest Studies in Philosophy*, 4, eds. P. French, T. Uehling H. and H. Wettstein, 73-121. Reprinted in *Externalism and Self-Knowledge*, eds. P. Ludlow and N. Martin, Stanford: CSLI, 21-84 (page references to the latter).

Cummins, R. (1983). *The Nature of Psychological Explanation*. Cambridge, MA: MIT Press.

Davidson, D. (1970). Mental Events. D. Davidson, *Essays on Actions and Events*, Oxford: Oxford University Press, 1980, 207-227.

Fodor, J. (1974). Special Sciences. J. Fodor, *RePresentations*. Cambridge, MA: MIT Press, 1981, 127-145.

Fodor, J. (1989). Making Mind Matter More. *Philosophical Topics*, 17 , 59-80.

Hardcastle, V. (1998). On the Matter of Minds and Mental Causation. *Philosophy and Phenomenological Research*, 58, 1-25.

Heil, J. and Mele, A. (1991). Mental causes. *American Philosophical Quarterly*, 28, 61-71.

Horgan, T. (1989). Mental Quausation. *Philosophical Perspectives*, 3, 46-76.

Jackson, F. and Pettit P. (1990). Program Explanation: A General Perspective. *Analysis*, 50, 107-117.

Kim, J. (1976). Events as Property Exemplifications. *Action Theory*, eds. M. Brand and D. Walton, Dordrecht: Reidel, 159-177. Reprinted in Kim 1993, 33-52 (page references to the latter).

Kim, J. (1988). Supervenience for Multiple Domains. *Philosophical Topics*, 16, 129-150. Reprinted in Kim 1993, 109-130.

Kim, J. (1992). Multiple Realization and the Metaphysics of Reduction. *Philosophy and Phenomenological Research*, 52, 1-26. Reprinted in Kim 1993, 309-335 (page references to the latter).

Kim, J. (1993). *Supervenience and Mind*. Cambridge: Cambridge University Press.

Kim, J. (1998). *Mind in a Physical World*. Cambridge, MA: MIT Press.

- Kim, J. (2005). *Physicalism, or Something Near Enough*. Princeton, NJ: Princeton University Press.
- Ladyman, J. and Ross, D. (2007). *Everything Must Go: Metaphysics Naturalized*. Oxford: Oxford University Press
- Macdonald, C. and Macdonald G. (2006). The Metaphysics of Mental Causation. *Journal of Philosophy*, 53, 539-577.
- Malcolm, N. (1968). The Conceivability of Mechanism. *Philosophical Review*, 77, 45-72.
- Marras, A. (1997). The Debate on Mental Causation: Davidson and His Critics. *Dialogue*, 36, 177-195
- Marras, A. (1998). Kim's Principle of Explanatory Exclusion. *Australasian Journal of Philosophy*, 76, 439-451.
- Marras, A. (2000). Critical Notice of *Mind in a Physical World*. *Canadian Journal of Philosophy*, 30, 137-160.
- Marras, A. (2007). Kim's Supervenience Argument and Nonreductive Physicalism. *Erkenntnis*, 66, 305-327.
- Orilia, F. (2008). Basic Ontology, Multiple Realizability, and Mental Causation. This volume.
- McLaughlin, B. (1989). Type Epiphenomenalism, Type Dualism, and the Causal Priority of the Physical. *Philosophical Perspectives*, 3, 109-136.
- Papineau, D. (1993). *Philosophical Naturalism*. Oxford: Blackwell.
- Papineau, D. (2002). *Thinking About Consciousness*. Oxford: Oxford University Press.
- Pearl, J. (2000). *Causality*. New York: Cambridge University Press.
- Richardson, R. (1979). Functionalism and Reductionism. *Philosophy of Science*, 46, 533-558.
- Shapiro, L. and Sober E. (forthcoming). Epiphenomenalism—the Do's and the Don't's. *Studies in Causality: Historical and Contemporary*, eds. G. Wolters and P. Machamer, Pittsburgh: University of Pittsburgh Press.
- Woodward, J. (1997). Explanation, Invariance, and Intervention. *PSA 1996*, vol. 2, 26-41.

Yablo, S. (1992). Mental Causation, *Philosophical Review*. 101, 245-280.